



The University of Bradford Institutional Repository

<http://bradscholars.brad.ac.uk>

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Access to the published online version may require a subscription.

Link to publisher's version: <https://doi.org/10.1007/s10796-019-09916-y>

Citation: Singh P, Dwivedi YK, Kahlon KS et al (2019) Smart monitoring and controlling of government policies using social media and cloud computing. Information Systems Frontiers. Accepted for Publication.

Copyright statement: © The Author(s) 2019. Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



Smart Monitoring and Controlling of Government Policies Using Social Media and Cloud Computing

Prabhsimran Singh¹ · Yogesh K. Dwivedi² · Karanjeet Singh Kahlon³ · Ravinder Singh Sawhney⁴ · Ali Abdallah Alalwan⁵ · Nripendra P. Rana⁶

© The Author(s) 2019

Abstract

The governments, nowadays, throughout the world are increasingly becoming dependent on public opinion regarding the framing and implementation of certain policies for the welfare of the general public. The role of social media is vital to this emerging trend. Traditionally, lack of public participation in various policy making decision used to be a major cause of concern particularly when formulating and evaluating such policies. However, the exponential rise in usage of social media platforms by general public has given the government a wider insight to overcome this long pending dilemma. Cloud-based e-governance is currently being realized due to IT infrastructure availability along with mindset changes of government advisors towards realizing the various policies in a best possible manner. This paper presents a pragmatic approach that combines the capabilities of both cloud computing and social media analytics towards efficient monitoring and controlling of governmental policies through public involvement. The proposed system has provided us some encouraging results, when tested for Goods and Services Tax (GST) implementation by Indian government and established that it can be successfully implemented for efficient policy making and implementation.

Keywords Cloud computing · E-government · GST · Sentiment analysis · Social media analytics · Twitter

1 Introduction

Traditionally, the general public had not been provided with opportunity to become a party to the public policy making decisions by the government agencies and mainly the bureaucrats used to have their own say (Severo et al. 2016). This led to a sharp decline in conviction as well as reliance of public towards the government in general and its policies in particular (Hibbing and Theiss-Morse 2001; James and Van Ryzin 2017; Rosenstone et al. 1993; Schneider et al. 2010). The perception

rift between the government and the public increased dramatically and currently both are struggling to maintain a perfect harmony with regard to their relationship (Janssen et al. 2018). Electronic-Government (E-Government) is one such powerful tool that holds all the essentials to improve the relationship between the general public and the government (Dwivedi et al. 2017a, b; Rana et al. 2016, 2017; Shareef et al. 2016) as it strongly put emphasis on maintaining transparency, enhancing public participation and upgrading the quality of service (Dwivedi et al. 2016). However, the required

✉ Prabhsimran Singh
prabh_singh32@yahoo.com

✉ Yogesh K. Dwivedi
y.k.dwivedi@swansea.ac.uk

Karanjeet Singh Kahlon
karankahlon@gndu.ac.in

Ravinder Singh Sawhney
sawhney.ece@gndu.ac.in

Ali Abdallah Alalwan
alwan_jo@bau.edu.jo

Nripendra P. Rana
n.p.rana@swansea.ac.uk

¹ Department of Computer Engineering & Technology, Guru Nanak Dev University, Amritsar, India

² Emerging Market Research Center (EMaRC), School of Management, Swansea University, Swansea, UK

³ Department of Computer Science, Guru Nanak Dev University, Amritsar, India

⁴ Department of Electronics Technology, Guru Nanak Dev University, Amritsar, India

⁵ Amman University College for Banking and Financial Sciences, Al-Balqa Applied University, Amman, Jordan

⁶ Emerging Market Research Center (EMaRC), School of Management, Swansea University, Swansea, UK

Information and Communication Technology (ICT) infrastructure, implementation and operational cost remained a major bottleneck towards the implementation of e-government in true sense. Cloud computing is one such popular as well as reliable technology that can provide an optimal solution of this delinquent issue (Mohammed et al. 2016). With several inbuilt advantages such as on-demand scalability and pay-as-use, have motivated many countries to adapt cloud computing based e-government (Sharma et al. 2012). However, the important point of public participation towards the policy making exercise still remains a distant dream. To overcome this paradox, governments all around world have started making use of social media to acquire appropriate feedback from the various realms of society regarding effectiveness of the new public policies being formulated (Grubmüller et al. 2013).

So, social media can overcome the limitations of e-government like low accessibility and public participation (Pieterse and Johnson 2011), thus improving the relationship between the general public and the government. But with more than 2 billion active users on social media (Aladwani and Dwivedi 2018), the volume of data generated is quite large (i.e. big data). The traditional systems are not sufficient to handle this much large volume of data, due to high infrastructural requirement to store and process this data. As discussed above cloud based e-government can be used to overcome this infrastructure related problems. Earlier studies (Cotterill and King 2007; Dwivedi et al. 2017; Kaigo and Okura 2016; Mossberger et al. 2013) have highlighted the role of social media in e-government but these studies failed to highlight the infrastructural barrier in its implementation.

This paper proposes an innovative and smart approach which unitizes the capabilities of two important concurrent technologies, cloud computing and social media analytics for efficient monitoring and controlling of government initiative for the general public. The main objective of this research is to envisage the effect of increased public participation towards policy making decisions. We have applied our proposed system to a latest policy implementation by the Government of India i.e. *Goods and Services Tax* (GST) which was intended to unify all the taxes in the country in a single domain and bring rationalization (Ahmad and Poddar 2009; Garima 2018; Jha 2018; Joseph et al. 2017). Earlier Indian Tax system consisted of complex structure where both state governments and central government imposed their own tax on a specific commodity. Due to the variable tax rates imposed by different state governments, same commodity was sold at different rates in different states. In order to overcome this complex problem, multiple solutions had been suggested from time to time. Eventually in 1999, a proposal was initiated by the council of ministers under the aegis of then Prime Minister of India, Mr. Atal Bihari Vajpayee (Singh et al. 2019) to standardized the tax structure throughout India. It took another 18 long years for GST to become a reality (Constitution 2017) and

finally, on July 1, 2017 GST was implemented pan India. Like every other government policy, it was the common man who faced the initial hardships. Misconceptions and confusion related to its implementation led to create a faulty image about the policy among the common people, although GST has many benefits in store for them (GST Benefits 2019). The initial hardships and the resolution provided by MoF motivated us to use this GST data for experimentation in our proposed system. The results have been observed to be quite encouraging, which establish the fact that the proposed approach can be instrumental in efficient evaluation of policy making decisions including future policy formulation.

The paper is structured into eight sections. The second section discusses brief literature regarding cloud based e-government and social media. Following that the third section provides a discussion on proposed methodology. Section four presents implementation and results. Section five gives performance analysis of the system, while section six discusses economic feasibility of the system along with limitation and future research directions. Finally, we conclude in the last section.

2 Review of Literature

The literature review is classified into the following two subsections. The first subsection highlights the use of cloud computing in e-government. While the second subsection highlights the use of social media for policy making.

2.1 Use of Cloud Computing in E-Government

Governments around the world are promoting multiple services to serve its citizens in a better and transparent way. E-Government is one such mechanism that enables the government to perform its day to day tasks and provide a better service to its citizens (Mohammed et al. 2016; Zwattendorfer et al. 2013). E-Government is the use of ICT and other web technologies to provide access to effective, efficient and transparent public services to its citizens and employees (Jeong 2006; Rana et al. 2013). Although, use of ICT can provide several advantages, yet the required technical infrastructure, implementation cost and requirement of skilled staff becomes major obstacle towards E-Government implementation (Rana et al. 2013). With the emergence of cloud computing, these challenges can be addressed up to a fair degree of satisfaction for all stakeholders (Mohammed et al. 2016). Cloud computing consists of large shared pool of computer resources which provide features like on-demand scalability and pay-as-use (Sadiku et al. 2014). These advantages have played a decisive factor in motivating the governments of many countries to migrate from traditional costly E-Government model to cost efficient as well as scalable cloud based E-Government model (Sharma et al. 2012). Due

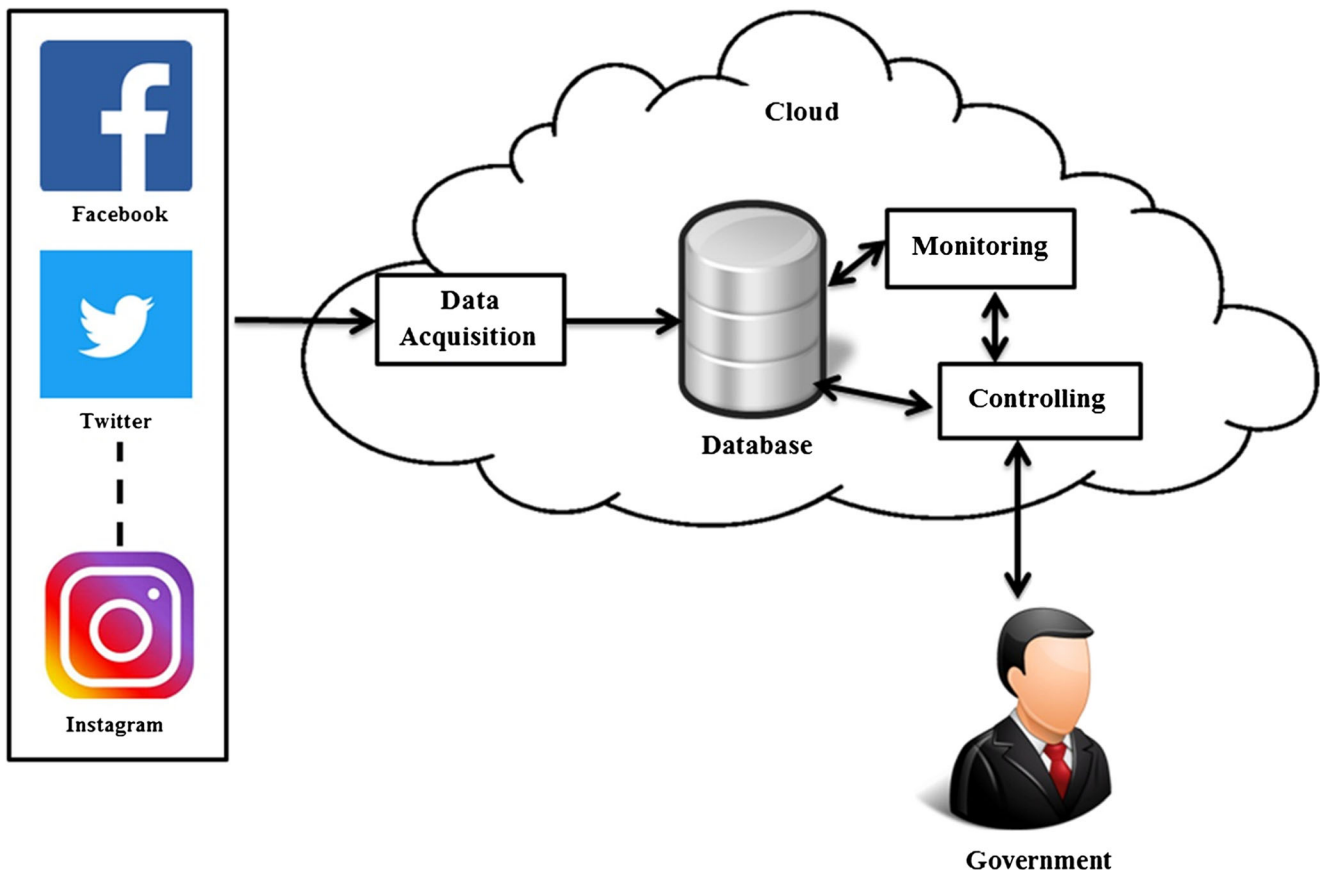


Fig. 1 Architecture of proposed system

to this adoption cloud computing has become a new channel for delivering improved government services (Liang et al. 2011; Smitha et al. 2012). Cloud based E-Government services provides advantages like reduced operational cost, distributed data storage, scalability and finally security management (Smitha et al. 2012), which indeed improves the relationship between government and public. Not only this, but cloud based E-Government model is also building a strong foundation for smart cities (Clohessy et al. 2014).

2.2 Use of Social Media for Policy Making

Traditionally, policy making had always been based upon official statistics (data) generated by government agencies and international bodies (Severo et al. 2016). However, decision makers normally reported flaws in this form of data because of publication delays, top down approach and insufficient topic of interest etc. (Pfeffermann et al. 2015). Therefore to overcome this traditional problem of data collection, political scientist and policy makers moved their attention towards social media platforms like Twitter and Facebook, to accumulate more authenticated user data and that too in real time. Hence, social media becomes an efficient tool for enhancing transparency in government working and increasing the communication between citizens and policy makers by bringing out

transparent data analytic statistics (Chung and Zeng 2016; Lee and Kwak 2012). Nowadays, social media have become an integral part of everyday life, irrespective of the status of any individual (AlAlwan et al. 2017; Dwivedi et al. 2015; Kapoor et al. 2018; Shiau et al. 2017, 2018). This virtual world provides a perfect platform for people from all around the world to discuss topics of common interest such as sports, entertainment and even politics. Talking about politics, at least 33% of social media users comments, discusses or post about politics on these platforms (Hossain et al. 2018). Even governments have realized the potential of social media. Consequently, various government agencies have started using various social media platforms to connect and engage with general public (Aladwani

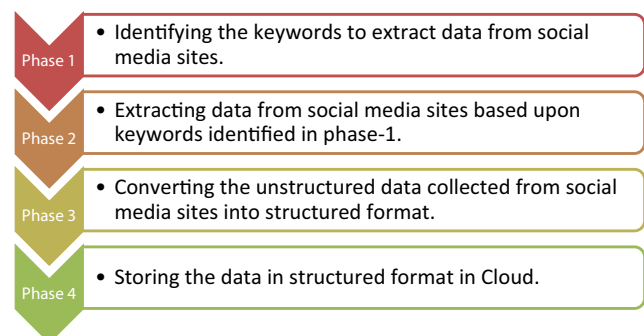
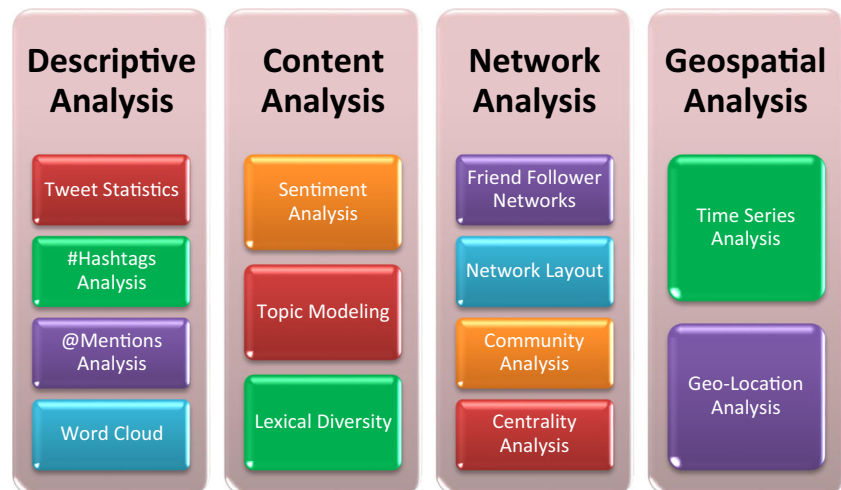


Fig. 2 Different phases of data acquisition component

Fig. 3 Different Social Media Analytics Techniques (Source: Adapted from Grover et al. 2017)



and Dwivedi 2018; Alryalat et al. 2017; Bertot et al. 2012; Rana et al. 2015). As social media is helping to increase the interaction between public and government, it is indeed facilitating public participation (Ceron and Negri 2016). Generally, people post something regarding government, politics or policies which might be intentional or unintentional (Hossain et al. 2018), which can be utilized by government for formulating more effective public policy as well as designing and delivering better services to its citizens (Androutsopoulou et al. 2018; Joseph et al. 2017; Park et al. 2016).

As represented in both the above subsections that considerable work has been done in the field of cloud computing based e-government and use of social media for policy making having their own advantages and benefits. Hence, cloud computing based e-government and social media can be considered as an influential paradigm to be collaborated with controlling and monitoring of public policies. With large number of people engaging on social media sites for discussions related to public policies, resulting in generation of huge

amount of data. There is a need for a cloud based system that can utilize this large data and map useful public opinion in form of issues, concerns, solutions, proposals, advantages and disadvantages (Androutsopoulou et al. 2018) towards a public policy at an early stage so that appropriate steps can be taken in order to please the public for which the policy has been formulated (Bertot et al. 2012). To the best of our knowledge, till date no effort has been made to combine both these services. Hence, this paper aims to unitize the capabilities of cloud computing and social media analytics for efficient monitoring and controlling of public policies. With this research, we try to answer the following questions that may erupt in the mind of the readers of this paper:

- Can social media act as an effective tool for opinion mining on issues related to national importance?
- Can cloud based environment be successfully employed in implementing a national policy in cost efficient manner?

Table 1 Benefits of cloud computing (Buyya et al. 2013, IBM 2019)

Features	Sub-Features	Benefits
Flexibility	Scalability	Cloud provides on demand scaling, depending on the requirement of the users.
	Storage options	User can choose from public, private or hybrid clouds for storage.
	Control choices	User can select which service to utilize i.e. software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS)
Efficiency	Tool selection	User can utilize prebuilt tools to build their application/system.
	Security features	Cloud saves the data in encrypted form to store it in a secure manner.
	Accessibility	User can access cloud application virtually from any device connected to internet
	Speed to market	Users can deploy their cloud application in a much faster manner.
	Data security	Hardware failure does not lead to data loss.
Strategic value	Savings on equipment	Since cloud uses remote resources, hence cost of resources is saved.
	Pay structure	Cloud offers pay per use facility, meaning we pay for whatever resources we use.
	Streamlined work	It is the responsibility of cloud service provider to take care of infrastructure, enabling the user to concentrate on application development.
	Regular updates	Cloud providers offer their user up to date technology.
	Collaboration	Since we can access the cloud through internet connected devices, the application development teams can make developments sitting in different parts of the world.
	Competitive edge	Cloud users have an edge over their rivals as they need not concentrate on resource management.

Table 2 Various attributes fetched during data collection

Sr. No.	Attribute	Details
1	Tweet ID	The ID of a particular tweet
2	Tweet	Actual tweet message
3	Sender	The person who tweeted the tweet
4	Date	Date on which tweet was tweeted
5	Time	Time on which tweet was tweeted
6	Place	Location from where tweet was tweeted

3 Research Methodology

The main aim of this study is to combine the capabilities of Cloud Computing and Social Media Analytics for efficient monitoring and controlling of public policies. For this cloud based system is proposed. Figure 1 shows the architecture of the proposed system. The proposed system comprises of three main components: (a) Data acquisition component (b) Monitoring component and (c) Controlling component. The reason to break-down the system into its three components was to separately highlight the task being performed by each component. The details of each component are explained as follow:

- (a) **Data acquisition component:** Data acquisition is the first component of our proposed system. This component is responsible for data collection from various social media platforms like Facebook, Twitter, Instagram etc. The data collection task is performed on continuous basis in a time specific manner. This component works in four phases. Figure 2 gives an overview of these phases. Phase-1 deals with identifying all the keywords (Search Terms) on basis of which data will be extracted from various social media sites. Once all the keywords have been identified, phase-2 comes into play. Phase-2 deals with extraction of data (Tweets, Posts etc) from various social media websites

(Facebook, Twitter, Instagram etc). Phase-2 returns data in unstructured json format, so in phase-3 this unstructured json data is converted into structured format (Excel Format). This excel file is stored on cloud database in phase-4 so that the computations can be effectively made. The details of each phase are explained in upcoming Section (See Section 5).

- (b) **Monitoring Component:** This component is responsible for performing the monitoring activities. These monitoring activities are performed using various social media analytics techniques, on the user content for extracting important decision making information (Stieglitz and Dang-Xuan 2013a). Figure 3 shows the various social media analytics techniques. Generally, social media analytics is categorized into four broad categories namely descriptive analysis, content analysis, network analysis and geospatial analysis (Singh et al. 2018a). The descriptive analysis gives descriptive statistics about number of tweets, number of tweet senders, hashtags, @mentions, word cloud etc. (Chae 2015; Ganis and Kohirkar 2015; McNaught and Lam 2010; Purohit et al. 2013). Content analysis deals with obtaining semantic content from the given text. This is achieved using techniques like sentiment analysis, topic modeling etc. (Kassarjian 1977; Kayser and Blind 2017; Llewellyn et al. 2015; Zhang et al. 2016). Network analysis allows us to identify different communities and cluster of users based upon their opinion towards an entity (Abascal-Mena et al. 2015; HerdaĖdelen et al. 2013; Stieglitz and Dang-Xuan 2013a, b). Finally, geospatial analysis deals with location based analysis and time series analysis (Attu and Terras 2017; Saboo et al. 2016; Singh et al. 2018b; Stephens and Poorthuis 2015). All these techniques give different results that are useful for appropriate decision making and these results are again stored in cloud database.

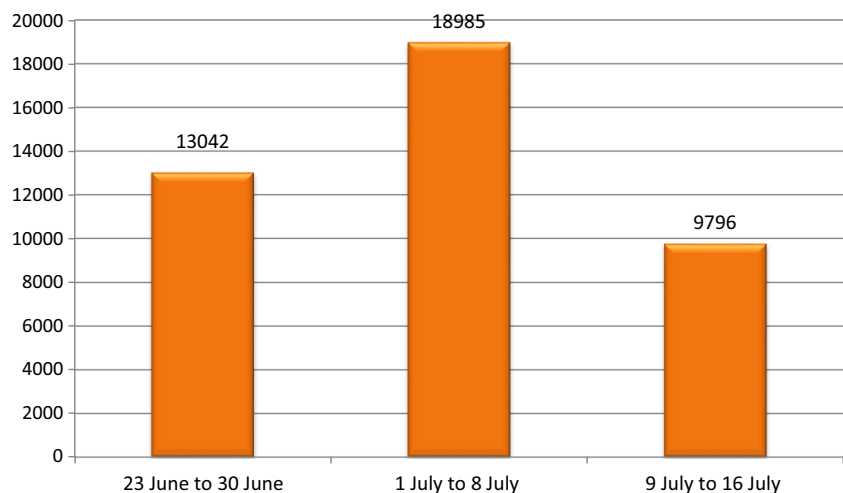
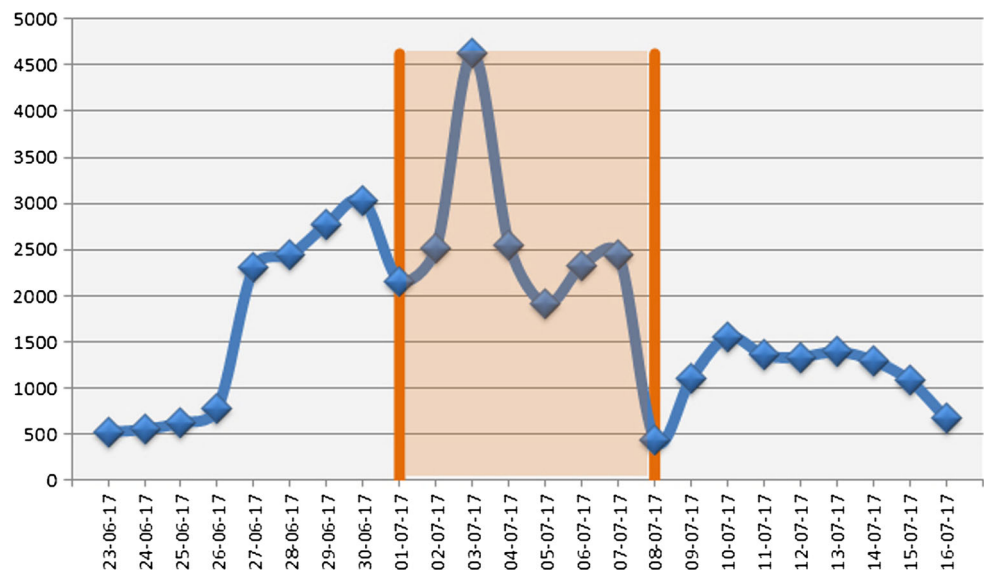
Fig. 4 Tweet collection details

Fig. 5 Daily tweet collection



- (c) **Controlling Component:** This component is responsible for decision making based upon the results generated in monitoring component and providing appropriate alerts to the government. The alerts are generated in a timely manner and the time interval can be fixed ranging from hours to days or even weeks and months. In addition to this it also sends calculated information to the government so that appropriate control measures can be taken based upon these results generated by the monitoring component.

4 Implementation and Results

Considering numerous advantages, the backbone of our proposed system is cloud. Though the system can be implemented without considering cloud computing, however the various features associated with cloud environment would make cloud based system more versatile and hence, motivate us to integrate cloud within our proposed system (Buyya et al. 2013; IBM 2019). The benefits associated with cloud computing are explained in Table 1.

Table 3 Tweet statistics

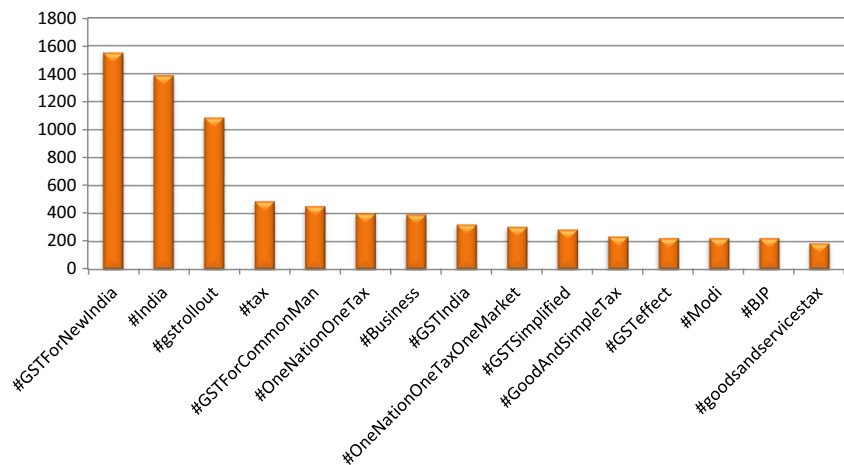
	Phase-1	Phase-2	Phase-3
Total Tweets	13042	18985	9796
Max Tweets in a Day	3026	4622	1552
Min Tweets in a Day	526	440	671
Total Unique Senders	11711	15187	8502
Average Tweets per Sender	1.11	1.25	1.15
Average Tweet per Phase	1630.25	2373.12	1224.5

For implementation on cloud we have used Amazon EC2 (AWS Amazon-EC2 2017), while for performing computations on cloud we have used RStudio server. There was no specific reason for selecting Amazon EC2 for implementation, in fact any alternate cloud service provider like Microsoft Azure (2018) or Google cloud (Google Cloud 2018) could have been used. The other integral part of our system is data from social media sites. All social media sites provide authorized APIs (Application Programming Interfaces) for fetching data from these sites in an authenticated manner. Now in order to fetch data from these social media sites, we need to identify keywords based upon which this entire search operation will be based. This section is divided into various sub-sections, each depicting a specific operation.

4.1 Data Collection

Data for experimentation has been collected from Twitter. Though the proposed system (See Section 3) is capable enough to use multiple social media platforms for data retrieval, however for this experimentation only Twitter data is used. As discussed above data was fetched based upon specific #hashtags (#GST, #GSTForNewIndia and #OneNationOneTax). For identification of #hashtags, an expert team consisting of three independent experts was constituted. These experts sorted #hashtags that were linked with GST. The common #hashtags identified by three experts were selected for performing tweet fetching operation (Singh et al. 2018c). The Twitter API used for fetching tweet provide us with various search parameters such as, language of tweets, type of tweets (Original, Re-Tweets or Both) and date range between which fetching operation need to be performed. Using these search parameters, we perform the fetching operation in more appropriate and desired manner.

In total 41,823 tweets were collected over a span of 24 days starting from June 23, 2017 to July 16, 2017. Since data was

Fig. 6 Top15 (#)Hashtags

fetches using APIs, it was in unstructured form i.e. json format. In order to use this data for further analysis it was necessary to convert this unstructured data into structured data i.e. excel format. Table 2 shows the various attributes that were fetched during data collected from Twitter API.

4.2 Tweet Statistics

In total 41,823 tweets were collected from 35,400 different users from India. In total 2873 users were detected who tweeted more than one tweet, accounting a total of 6423 tweets. For better interpretation of the results tweet

collection period was broken into 3-phases (Pre-GST, In-GST and Post-GST) of eight days each. The breakup of these phases is shown in Fig. 4, while Fig. 5 shows daily tweet collection.

The shaded part (orange color) in Fig. 5 indicates the In-GST period when Twitter traffic was highest as compared to other two phases. The details of tweet statistics are given in Table 3.

Inference: These statistics helps the government in understanding the scale and impact of the issues. The data clearly indicates that the social engagement created by GST was on

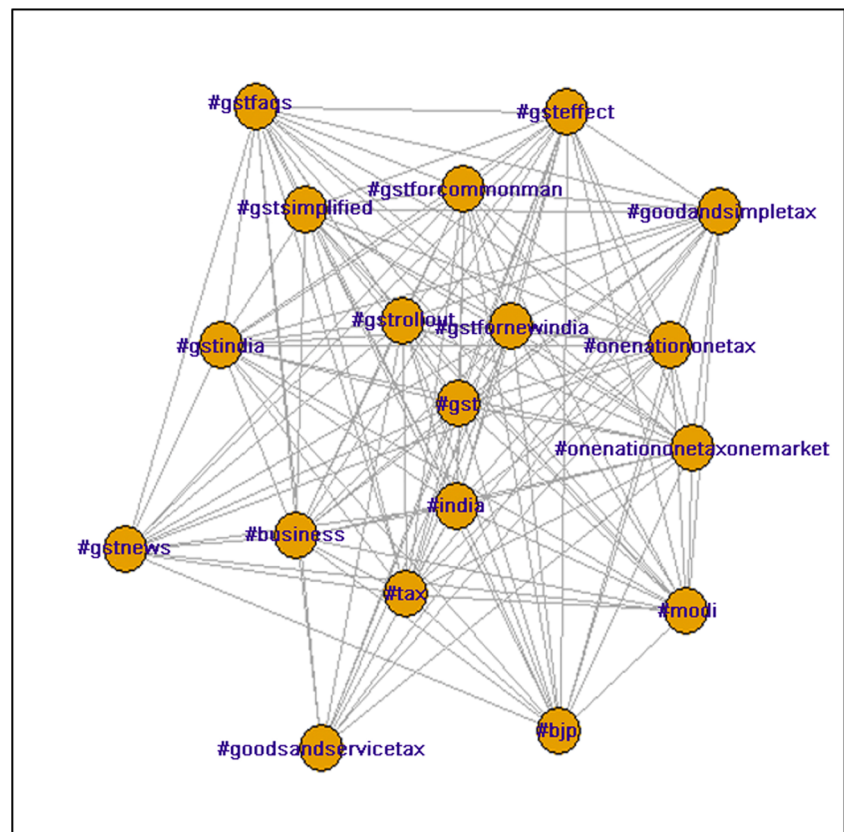
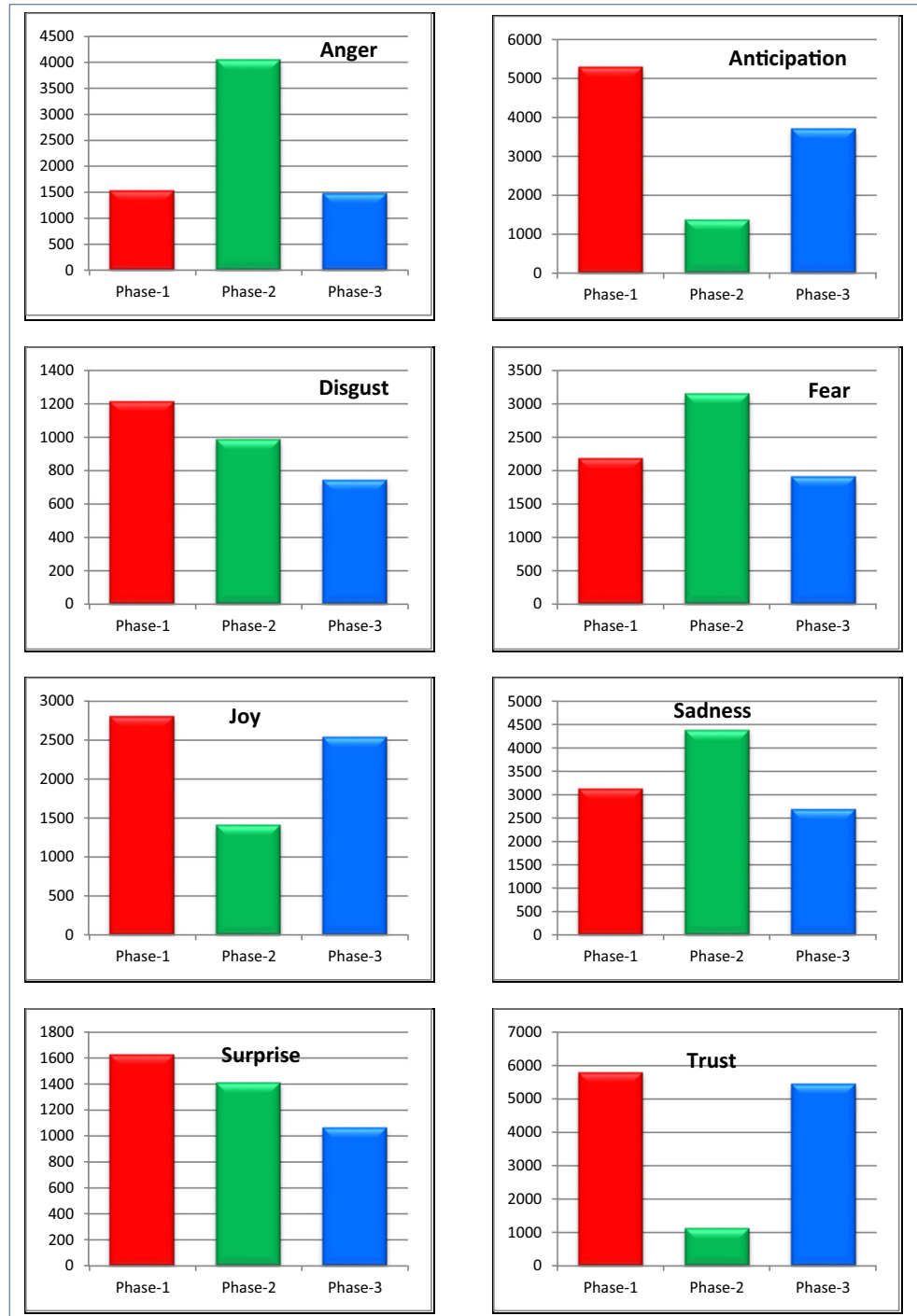
Fig. 7 Association among frequent hashtags

Table 4 Hashtag adjacency matrix

	#gst	#gsteffect	#gstforcommonman	#gstfornewindia	#gstrollout	#tax	#onenationonetaxonemarket	#gstindia	#gstsimplified	#india
#gst	18,925	221	451	1536	1075	485	304	323	284	1329
#gsteffect	221	225	26	90	62	3	2	8	6	10
#gstforcommonman	451	26	453	164	52	11	2	5	27	14
#gstfornewindia	1536	90	164	1549	312	18	71	35	83	63
#gstrollout	1075	62	52	312	1088	21	38	53	64	58
#tax	485	3	11	18	21	489	1	7	19	103
#onenationonetaxonemarket	304	2	2	71	38	1	305	5	12	6
#gstindia	323	8	5	35	53	7	5	323	29	11
#gstsimplified	284	6	27	83	64	19	12	29	285	15
#india	1329	10	14	63	58	103	6	11	15	1338
#onenationonetax	396	7	8	71	30	4	24	6	4	18
#goodandsimpletax	229	9	7	54	30	1	7	2	2	4
#gstnews	219	0	0	1	9	8	0	0	2	20
#business	390	2	1	9	10	11	1	1	3	247
#modi	222	3	3	22	14	11	3	4	2	40
#bjp	218	3	1	11	10	6	3	5	1	22
#gstfaqs	175	4	7	15	64	5	9	25	57	8
#goodsandservicestax	188	2	0	10	7	0	4	0	0	2

	#onerationonetax	#goodandsimpletax	#gstnews	#business	#modi	#bjp	#gstfaqs	#goodsandservicestax
#gst	396	229	219	390	222	218	175	188
#gsteffect	7	9	0	2	3	3	4	2
#gstforcommonman	8	7	0	1	3	1	7	0
#gstfornewindia	71	54	1	9	22	11	15	10
#gstrollout	30	30	9	10	14	10	64	7
#tax	4	1	8	11	11	6	5	0
#onenationonetaxonemarket	24	7	0	1	3	3	9	4
#gstindia	6	2	0	1	4	5	25	0
#gstsimplified	4	2	2	3	2	1	57	0
#india	18	4	20	247	40	22	8	2
#onenationonetax	402	16	0	0	6	3	2	3
#goodandsimpletax	16	235	0	2	6	1	2	4
#gstnews	0	0	222	1	0	0	1	0
#business	0	2	1	393	2	1	3	1
#modi	6	6	0	2	225	28	2	0
#bjp	3	1	0	1	28	224	0	0
#gstfaqs	2	2	1	3	2	0	176	3
#goodsandservicestax	3	4	0	1	0	0	3	188

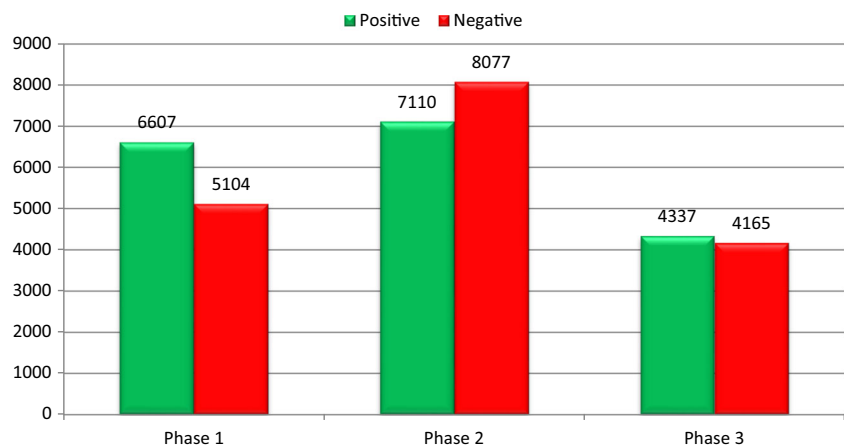
Fig. 8 Results of E-Motion analysis

larger scale. As evident from the data the number of tweets increased by 45% from phase-1 to phase-2, depicting the issue has increased and larger number of people joined the conversation to put forward their reviews regarding this policy. This indeed acts as an indicator for the government to look into such scenarios and act accordingly.

Similarly the average tweets per sender are 1.11, 1.25 and 1.15 for the 3 phases respectively. This indicates that data collected is normalized (equally distributed) and not biased.

In case average tweets per sender is more, then there is a possibility that the data is biased and hence we need to normalize this. We can normalize the data by using a similar technique as used by Bornmann and Haunschild (2016). We can also restrict the number of tweets from a sender to one tweet per user by adopting a similar technique used by Singh et al. (2017). Since average tweets per sender are almost “1”, hence both the above discussed techniques are not required in these experimentations.

Fig. 9 Results of polarity analysis



4.3 (#)Hashtag Analysis

Hashtag analysis deals with the various hashtags that occurs among the collected tweets. In total 11,499 unique hashtags were found, that appeared 81,054 times. From the total 41,823 tweets, 19,114 tweets contained more than one hashtag. The analysis shows that hashtag “#GST” had the maximum occurrences i.e. 18,925. Figure 6 shows top 15 hashtags which had maximum occurrences. For better visualization “#GST” was excluded from the results in Fig. 6. Figure 7 shows association among popular hashtags. The association is plotted using Fruchterman-Reingold layout (Fruchterman and Reingold 1991; Csardi and Nepusz 2006). Similarly Table 3 shows the hashtag adjacency matrix. Both Fig. 7 and Table 4 the relationship among different hashtags.

Inference: Hashtags allows the government to understand the latest trends among the people. Any new policy will lead to

increase in people participation in discussion on Twitter (social media). This discussion can be in favour as well as against the new policy. So, hashtag analysis enables the government to keep a check on any trends generated among general public and taking appropriate measures to resolve the issues. Hashtag analysis shows that overall people were in support of GST, this is evident from the fact that our analysis shows most popular hashtags were “#gstforcommonman”, “#gstindia”, “#gstfornewindia”, “#onenationonetaxonemarket” etc.

However for a brief period during phase-2 (In-GST Period), hashtags such as “#Confussion”, “#GSTConfusion” and “#Clarification” were trending, indicating that people were not clear with GST rules and were raising the voice using these hashtags on twitter. But as soon as government took preventive measures like introduction of online course for GST training (NDTV News 2017) and restricting the retail sellers to charge more the MRP (Business Standard News 2017a, b), the trends were again in favour of GST.

Fig. 10 Number of positive tweets (day wise)

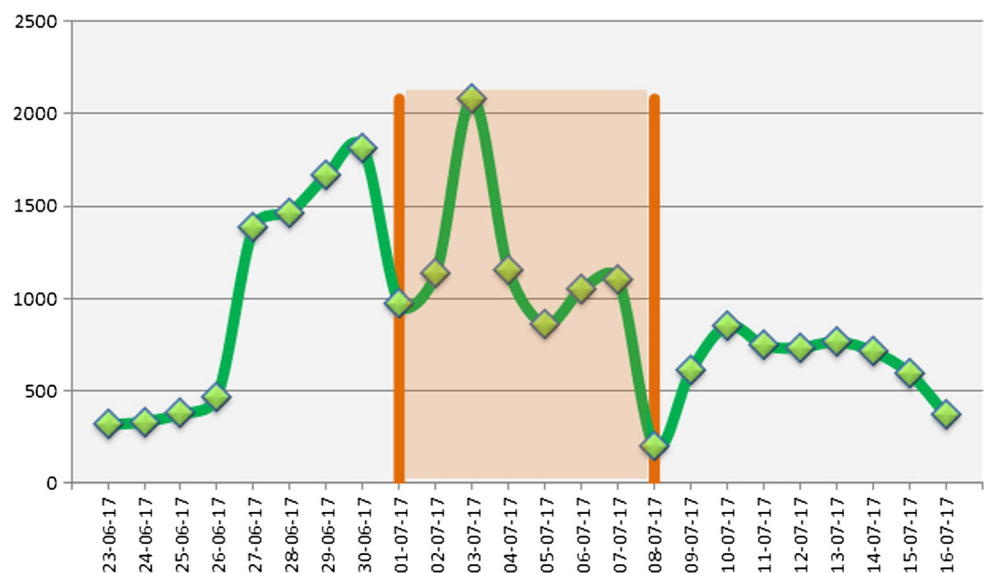
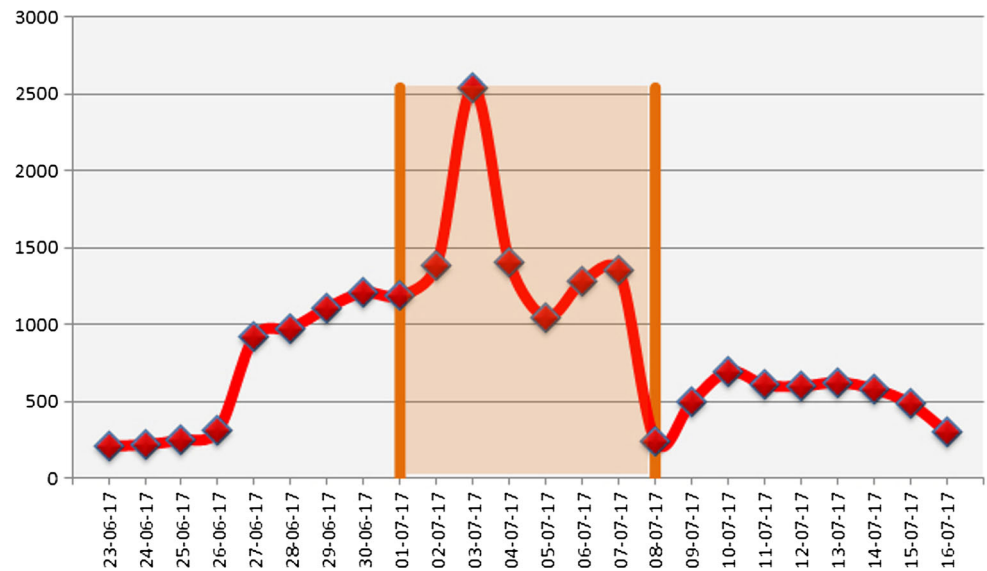


Fig. 11 Number of negative tweets (day wise)



4.4 Sentiment Analysis

Sentiment analysis is often regarded as most effective tool to map public response towards an entity (Joseph et al. 2017). Sentiment analysis is defined as a text analytic technique which deals with extraction of sentiment from given piece of text (Liu 2012; Mishra and Singh 2016). Unlike the previous analysis, the upcoming analysis requires preprocessing of data. Data preprocessing is an important task, which aims to prepare data for various data mining task (Van Broeck et al. 2005; García et al. 2016; Liu et al. 2015). Since Twitter data contained lot of noise and unwanted stuff, task of data preprocessing was performed (Haddi et al. 2013). These tweets

preprocessed by performing various task conversion to lower case, removing punctuations, removing special characters and finally removing web links. Sentiment analysis consists of two sub-operations: (a) E-Motion Analysis (Mohammad and Turney 2010; Ou et al. 2014) and (b) Polarity Analysis (Saif et al. 2013; Yuan et al. 2017).

- (a) **E-Motion Analysis:** It is a sentiment analysis operation, in which given data is classified according to emotion lexicon comprising of words having association with eight emotions (Trust, Surprise, Sadness, Joy, Fear, Disgust, Anticipation and Anger). The result of emotion analysis is shown in Fig. 8.

Fig. 12 Treemap of positive and negative words



Table 5 Classification of warning signals

Threshold Value	Warning Type	Interpretation
$\mu < 50\%$	Normal Situation	People are happy with the new policy
$50\% > \mu$ and $\mu < 65\%$	Primary Warning	A small section of people are not happy with the new policy
$65\% > \mu$ and $\mu < 80\%$	Secondary Warning	A considerable section of people are not happy with the new policy
$\mu > 80\%$	Tertiary Warning	People are not happy with the new policy

Inference: Phase-1 (Pre-GST Period) dominates emotions such as trust, anticipation and joy indicating support of people for GST. These results totally got reversed during phase-2 (In-GST Period) as emotions like fear and sadness dominated, showing unrest among citizens and alerting government to take appropriate steps. As soon as appropriate steps have been taken by government results of phase-3(Post-GST Period) started to fall in line with phase-1 (Pre-GST Period), indicating support for GST once again.

- (b) **Polarity Analysis:** Polarity analysis deals with polarity identification i.e. positive or negative. The result of polarity analysis is shown in Fig. 9, while the daily number of positive tweets is shown in Figs. 10 and 11 shows daily number of negative tweets. The shaded portion (Orange color) in Figs. 10 and 11 shows the In-GST period. Similarly, treemap of positive and negative words is shown in Fig. 12.

For better interpretation of the results we use a threshold value (μ) as given in Eq. 1. This threshold value will act as an indicator for the government act when the situation starts becoming out of control i.e. opinion of people towards the new policy is on the negative side.

$$\mu = \frac{\sum \text{Negative Tweets}}{\sum \text{Tweets}} \quad (1)$$

Table 6 Coherence score for different number of topic

Number of topics	Coherence score
2	0.4106
5	0.3481
8	0.3611
11	0.3456
14	0.4483
17	0.4449
20	0.4626
23	0.4814
26	0.4897
29	0.5082

For our calculations we have taken $\mu = 50\%$ (Threshold Value) i.e. whenever $\mu > 50\%$, this will be alarming signal for the government that citizens are not happy with the policy and some appropriate measure are required to overcome this unrest. Further, the warning signals have been further classified as primary, secondary and tertiary. The details of the same are shown in Table 5. Depending upon the type of warning the government needs to take appropriate steps.

Inference: During phase-1 (Pre-GST Period) the overall threshold value $\mu < 50\%$, this shows that citizens were in favour of GST as it ended multi tax system and hence giving a hope that prices of fast moving consumer goods (FMCG) will go down (Economic Times 2017). However, as soon as GST was actually implemented on July 1, 2017 (In-GST Period) the threshold value $\mu > 50\%$, showing unrest among citizens. This unrest was due to various problems encountered by public as soon as GST was implemented. From consumer's perspective negative sentiment was due to the fact that traders were charging GST over and above maximum retail price (MRP). Similarly, the traders were unhappy because of the confusion about GST rates on various goods plus they needed to update their current inventory system which led to increased expenses. Since we discussed that as soon as threshold value $\mu > 50\%$, it is an alarming signal for the government to take appropriate steps. So acting upon this, government soon issued warnings to traders that they cannot charge over MRP for any good (Business Standard News 2017a, b). In addition

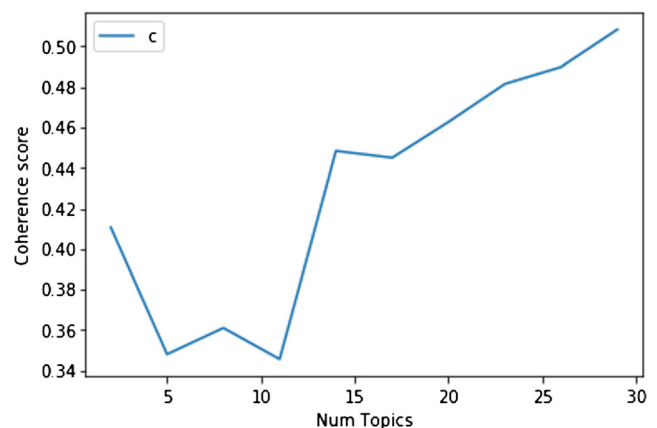
**Fig. 13** Coherence score for different number of topics

Fig. 14 Intertopic distance map



to this government also introduced various courses to train traders and make them familiar with GST filling and other related processes (NDTV News 2017). As a result of these measures by the government the threshold value again became $\mu < 50\%$ during phase-3 (Post-GST). This provided a concrete proof that the actions taken by government after phase-2 did helped to win the trust of citizens.

4.5 Topic Modeling

Topic modelling identifies the main themes among the captured tweets (Blei 2012). It is defined by the common set of words that have high probability of belonging to a particular topic. For performing the process of topic modelling latent dirichlet allocation (LDA) is used (Arun et al. 2010; Deveaud et al. 2014). For deciding optimum number of topics

we calculated the coherence score of topics (2 to 30). The results of these are shown in Table 6, while the scores are shown pictorially in Fig. 13. The results show that 29 topics are optimum. The results of topic modelling are shown in Fig. 14, in form of intertopic distance map.

Inference: The intertopic distance map helps to better understand the results. The size of circle represents the amounts of tweets associated with the topic. The size of topic-1 is the largest, indicating large portion of positive tweets. Though there were 29 different topics, however topic-1 (representing positive tweets), topic-2 (representing negative tweets) were visible as distinct one. Rest all (3 to 29) are all over lapping each other, indicating that all these topics are related to tweets where people have some sort of confusion or are demanding some sort of clarification on GST implementation. In order to

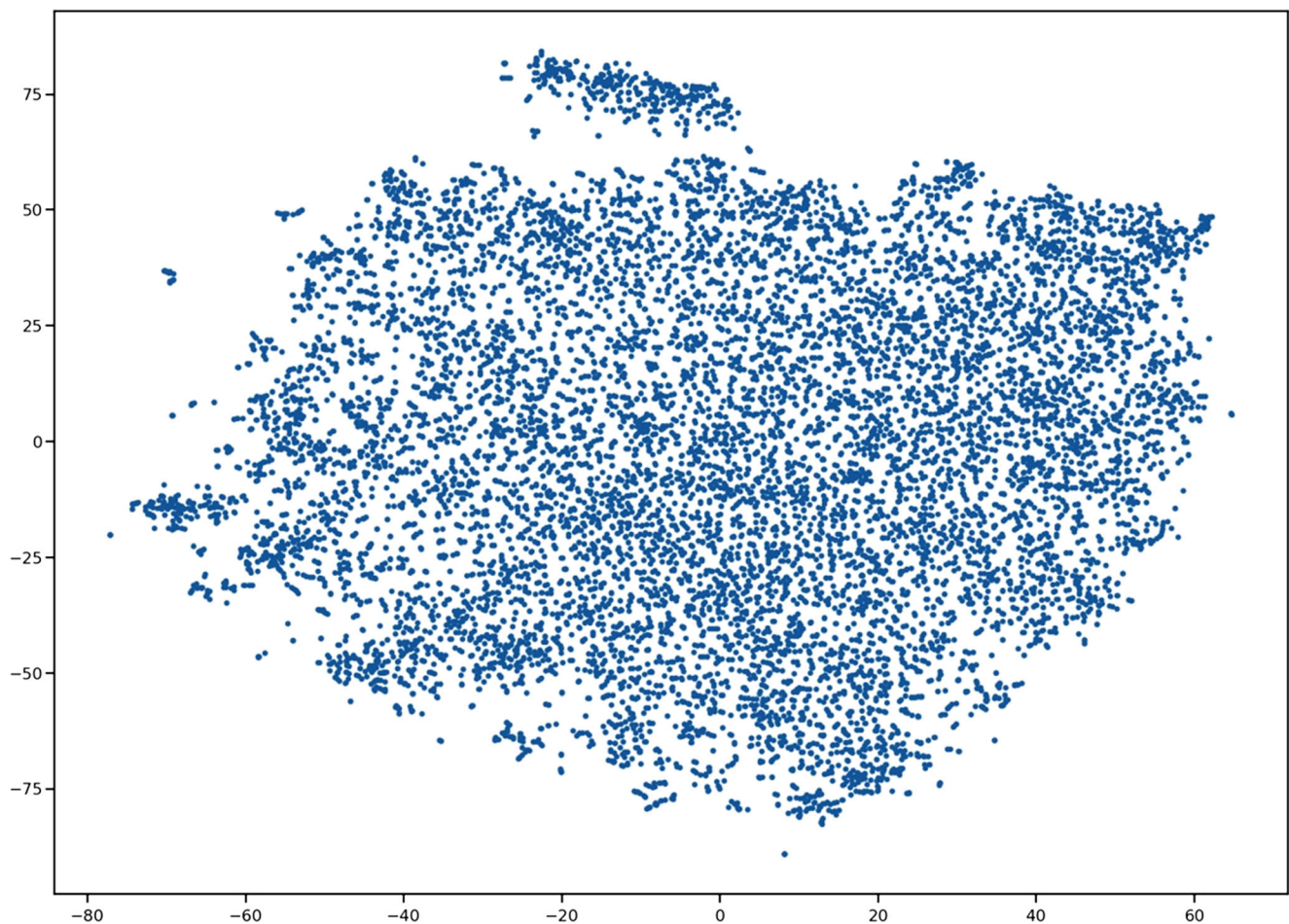


Fig. 15 Results of cosine similarity

get better interpretation of these results we calculated cosine similarity (Dehak et al. 2010) of tweets in topics 3 to 29. The results of cosine similarity are shown in Fig. 15. The results show moderate relationship among words of tweets in topic 3 to 29. If we zoom at particular portion as shown in Fig. 16, we can see words like “clarification”, “filling”, “gstreturn” etc. indicating that people wanted clarification in GST filling. So topic modelling can help the government to understand the main topics (themes) that are prominent among people leading to engagement on social media (Twitter). As indicated from results here the people wanted some clarification regarding GST filling and GST registration.

4.6 Community Analysis

Community analysis is part of network analysis, which detects various communities taking active part in a particular discussion (Ding 2011). The polarity (positive and negative) wise results of community analysis is shown in Fig. 17 for all three phases. In this the green color nodes indicates community which is happy with GST implementation, while red color

nodes indicates the community which has negative opinion towards GST. The results clearly indicates that the number of people with negative opinion towards GST increased during phase-2 (In-GST Period). But once all preventive steps were taken positive community again had upper hand.

The overall results of community analysis is shown in Fig. 18. The results shows that there are four prominent communities. Community-1 (Blue) indicates people who are supporting GST, making it the largest community. Community-2 (Red) indicates people which are against GST. Finally, community-3 (Green) and community-4 (Yellow) showing people who have some confusion and requires some clarification regarding GST.

Inference: The community analysis helps the government to identify prominent communities in order to check whether the newly implemented policy forms large positive community. Further, it can also assist the government to identify communities that have some sort of confusion or problem regarding the newly implemented policy, so that appropriate steps can be taken.

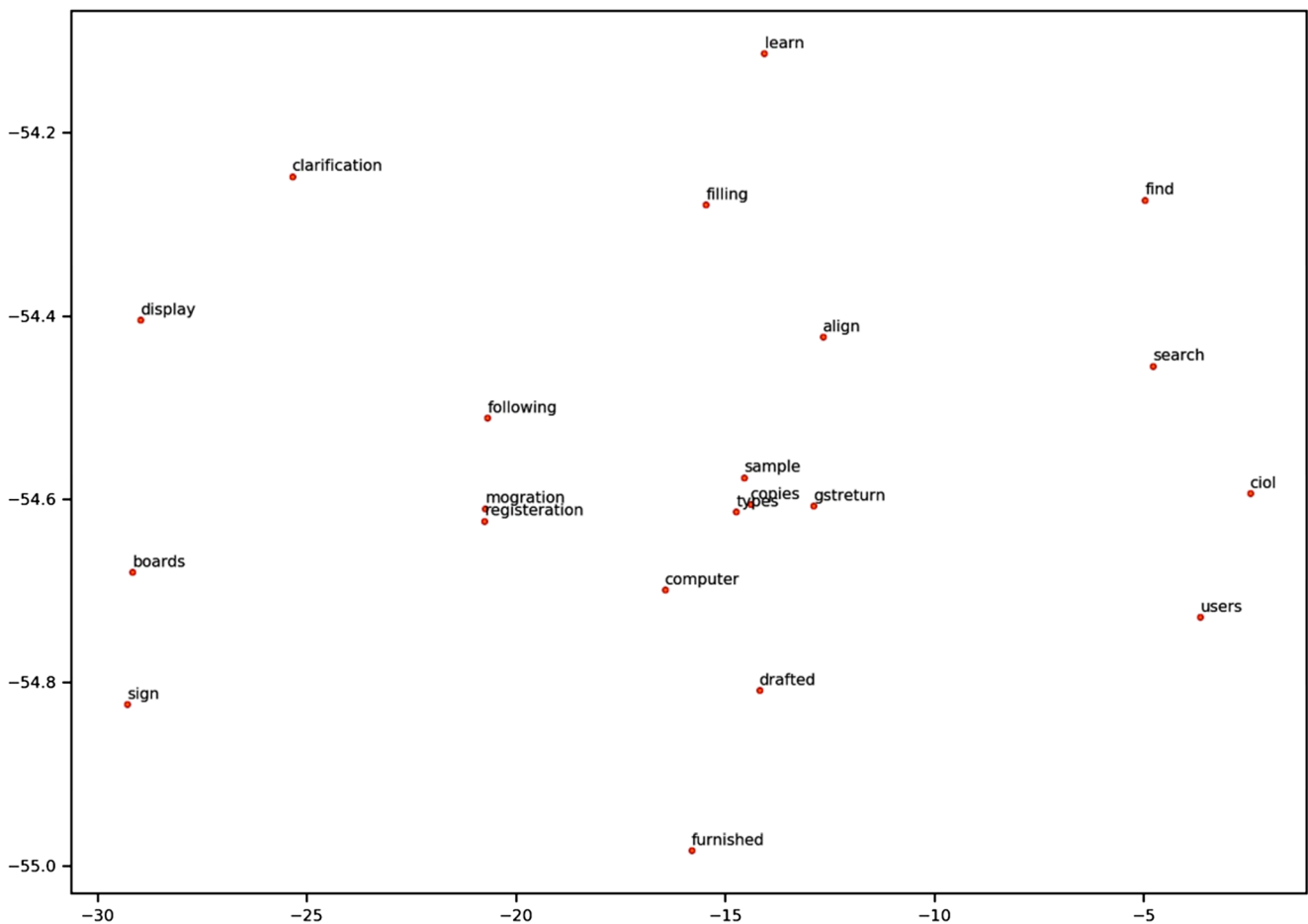


Fig. 16 Zoomed results of a particular portion

4.7 Geo-Location Analysis

Location based analysis is a very crucial tool for gathering information, while mapping public response towards an entity (Amirkhanyan and Meinel 2017). Although all tweets do not contain location from where they were tweeted, yet we can't

ignore them as they provide important information about the actual location, hence helping policy makers to target the audience while finding solution to their problem. Since in the previous section we detected that during phase-2 (In-GST Period) the overall sentiment was on negative side, it is essential to see which states and cities are worst affected so that appropriate

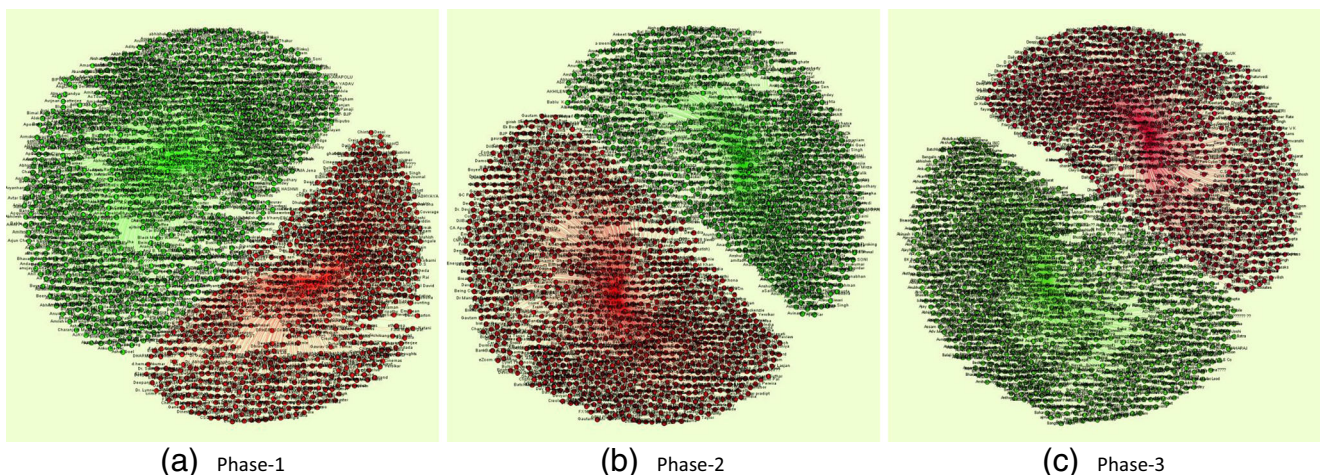
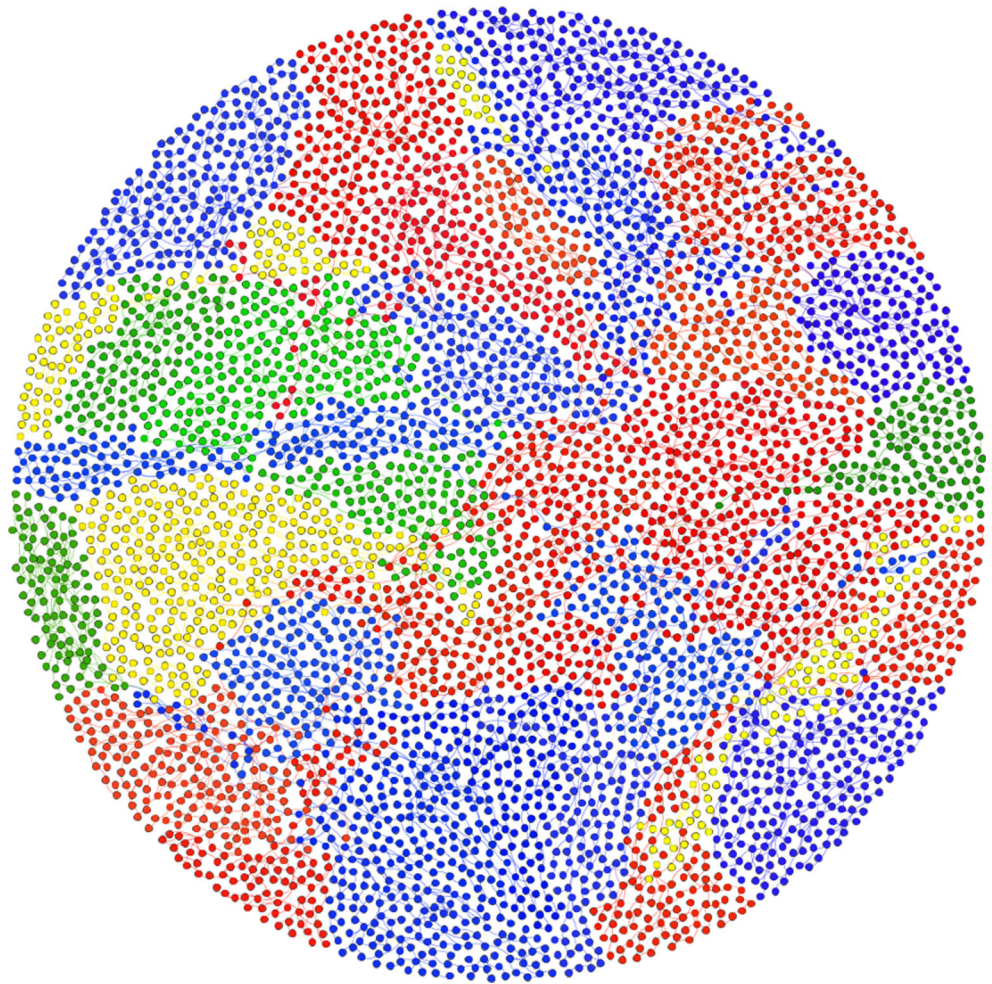


Fig. 17 Phase wise results (polarity) of community analysis

Fig. 18 Overall results of community analysis



solution can be found. Figure 19 shows results of zone wise analysis, while Fig. 20 shows results of state wise analysis and finally Fig. 21 shows location wise analysis of two markets of Mumbai city from where negative tweets were sent.

Inference: This location based analysis is extremely important; as it gives us the targeted audience which is unhappy with the policy and hence government can take appropriate steps keeping in mind the demands or problems of this targeted group. Since these two points indicates markets of Mumbai city, hence it shows that traders of these markets are not happy and government needs to take appropriate steps accordingly.

5 Performance Analysis

Since we are proposing this system for real time implementation, therefore it is very important to check the performance of system in a cloud based environment. As discussed earlier the entire experimentation has been performed on Amazon EC2

using compute optimised c4:large instance (containing 2 processors and 8GB of RAM). Since topic modelling is computationally very expensive (Grover et al. 2017), so the entire experimentation of performance analysis has been performed on topic modelling only. The collected data was broken up and the performance analysis was computed for 5000 tweets going all the way to 40,000 tweets, incrementing 5000 tweets in next iteration every time. The results of performance analysis are shown in Figs. 22 and 23 for resource utilization and execution time respectively. The results clearly indicate that resource utilization increases with increase in number of tweets (data). Similarly, execution time increases with increase in number of tweets (data).

6 Discussion

Methodologically, this study presents a cloud based system in which user generated social media content (tweets) is collected; and how this data can be utilized in efficient monitoring

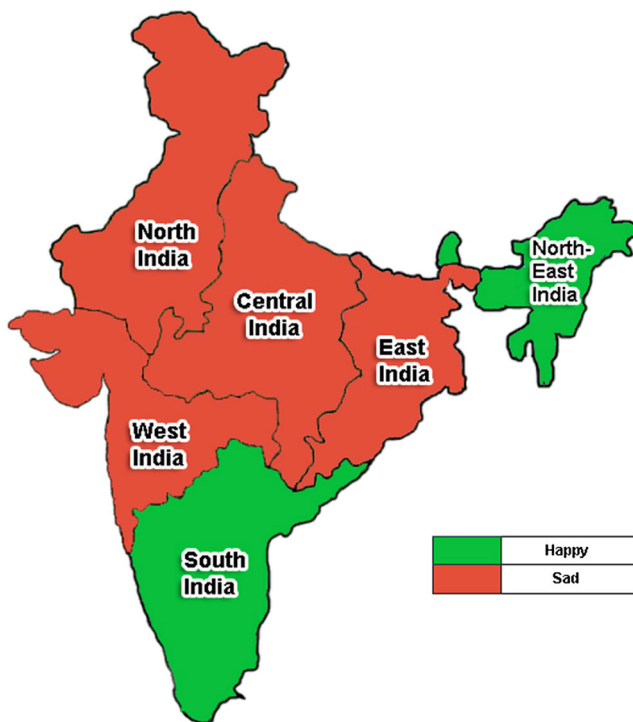


Fig. 19 Results of zone wise analysis

and controlling of public policies using social media analytics like descriptive analysis, content analysis, network analysis and geospatial analysis.

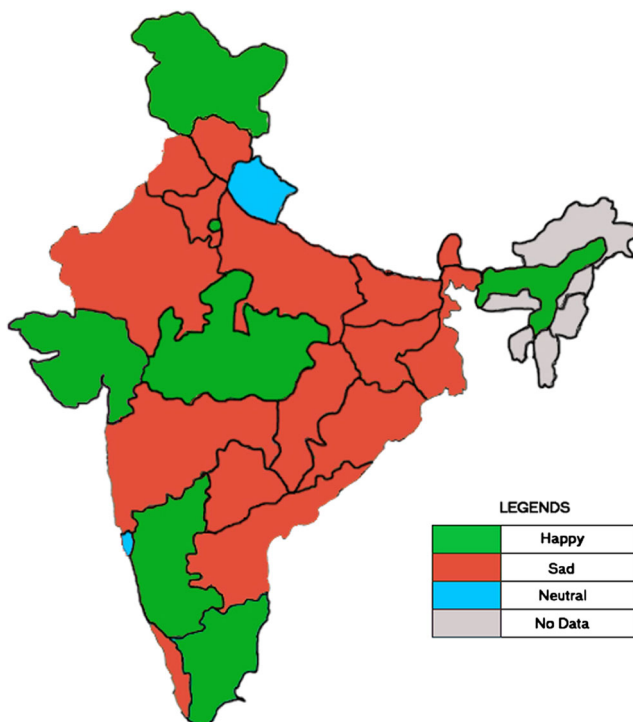


Fig. 20 Results of state wise analysis

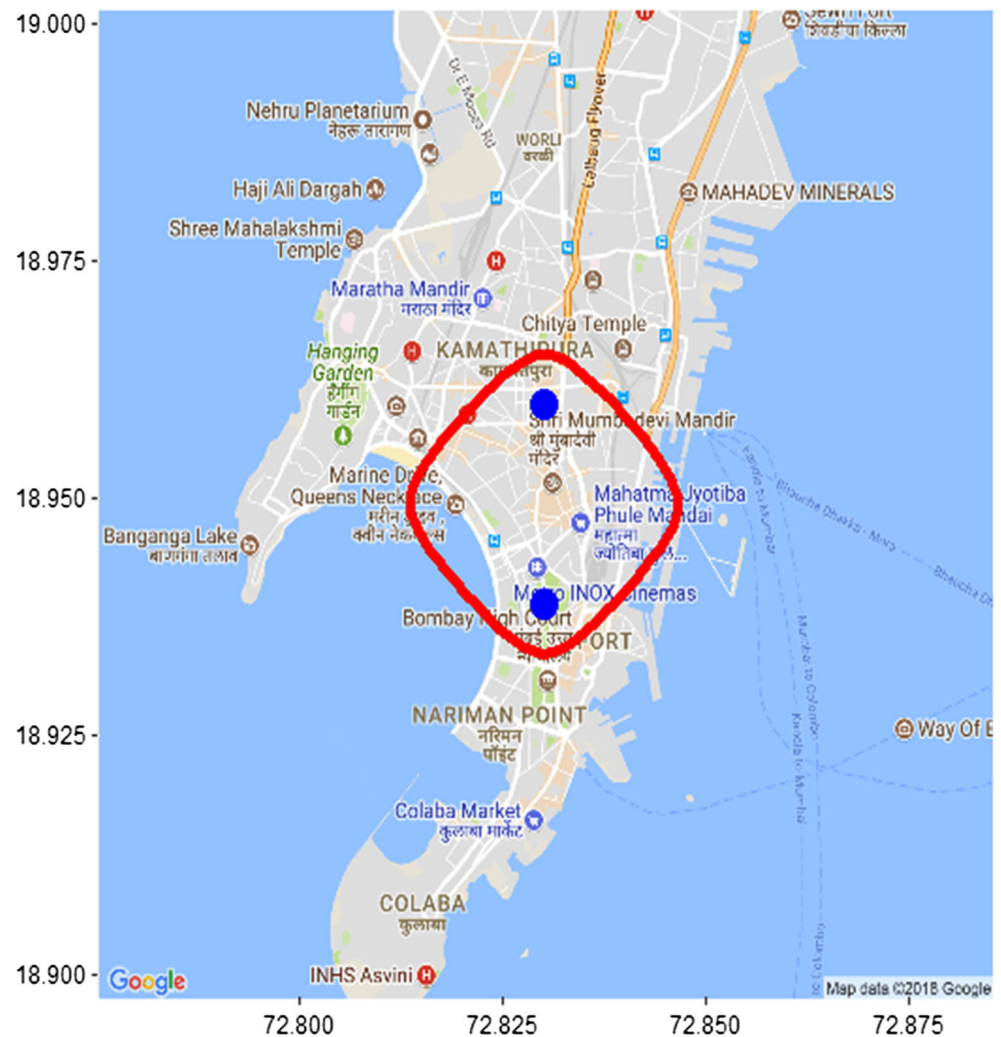
This study tries to map public opinion towards GST implementation by the Indian Government. The data was collected from Twitter using hashtag “#GST”. Different social media analytics were applied in cloud based environment (RStudio Server on Amazon EC2) to have an insight of GST policy from public perspective. The study of tweet statistics (Purohit et al. 2013) helped in understanding the impact and scale of the issue caused by implementation of the new policy. The trends were detected using (#)hashtag analysis (Chae 2015). Not only this, it also helped in understanding association between various hashtags. Sentiment analysis (Saif et al. 2013), gave polarity and emotions associated with tweets, which helped in understanding the opinion of people toward the GST. The sentiment analysis also served as an indicator for calculating threshold value, based upon which warning signals were generated for the government (See Section 4.4 Sub-section (b)). The identification of theme was done by topic modeling (Llewellyn et al. 2015). Network analysis (Walther and Kaiser 2013) helped in detection of prominent communities, engaged in GST discussion. Finally, geo-location analysis (Singh et al. 2018b) helped in detection of targeted audience which had negative opinion towards GST. The threshold based warning signals and location based analysis through which we can map the unhappy target audience up to market level were the unique contributions made to this study.

The results clearly indicate that in the starting (phase-1) there was a strong support for GST, but as soon as GST became reality (phase-2) and people started facing hardship the support for GST dipped and negative opinion rose among the people. Observing the dip in popularity of GST, the government acted fast and took steps to ensure that all confusion and hardship faced by public can be resolved. As soon as these preventive measures were taken, the overall opinion again changed to positive in the final phase (phase-3).

6.1 Economic Feasibility Analysis

The main reason for integrating cloud computing in our proposed system was cost cutting, as we only pay for what we use and initial cost of installation and later cost of maintenance are all eliminated. Since cost plays an important role in implementation of such projects and often becomes a major bottleneck in developing countries (Like India). Hence it becomes extremely important to evaluate economic feasibility of our proposed system. The very basic instance (t2.nano with 1 processor and 0.5GB RAM) available at Amazon cost around \$0.0062 per hour (Amazon EC2 Instance Price 2018). While the instance used by us for the entire experimentation c4.large instance (containing 2 processors and 8GB of RAM) cost around \$0.10 per hour. Depending upon the hardware requirements for computation we can switch to any instance. Another cost

Fig. 21 Results of location based analysis



that needs to be considered here is social media sites API cost. Again depending upon the use the government might opt for monthly or annual contract (Twitter API Price 2018). Paid APIs

gives access to various additional parameter like location of all tweets, which is not available in free APIs. So from above we can conclude that the proposed system is highly cost effective.

Fig. 22 Results of resource utilization

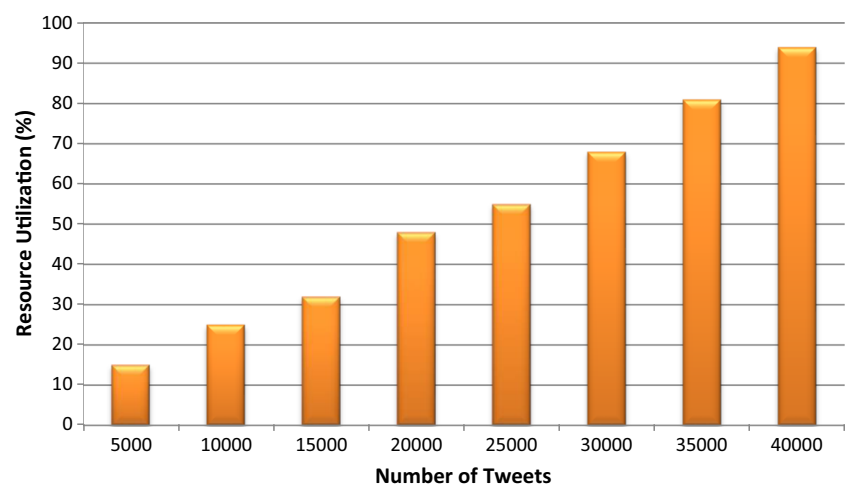
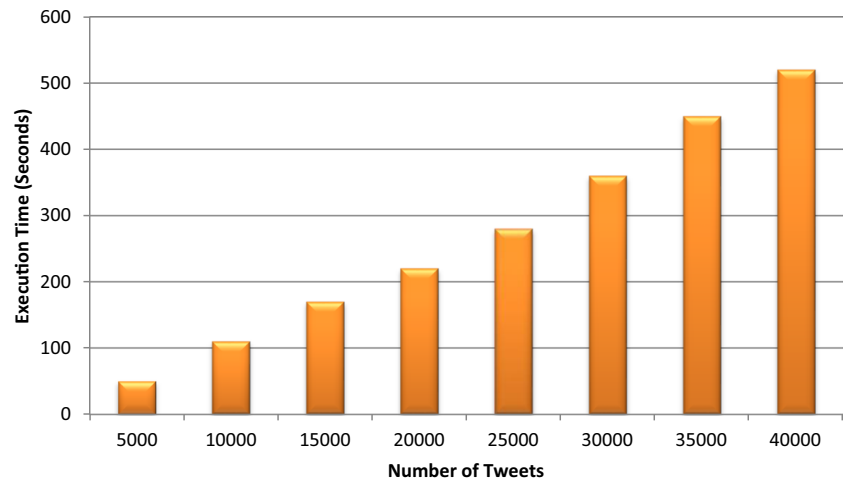


Fig. 23 Results of execution time (seconds)



6.2 Implications for Practice and Policy

The implication of the study for practice can be divided into three subsections (a) Government which will be implementing any policy (b) General Public for which this policy will be implemented and (c) cloud service providers. These subsections are discussed briefly below.

- (a) **Government:** The government is one of the main stakeholders in this system, as it will be the one who will be implementing the policy for the public and then monitor it using the proposed cloud based system. The first and foremost thing the government needs to do is to spread awareness among people before implementing any policy so that people passively (Androutsopoulou et al. 2018) start discussing about the new policy on social media platforms. Since the data is collected from social media platforms, and if people are not discussing so there will be no data available for the government to monitor.

For this government can create a page on Facebook, or highlight the advantages of new policy through tweets on Twitter. Early discussion on social media platform will help the government to better monitor the policy before actual implementation by considering various suggestions highlighted by people on social media. Another important insight that the government can gain are, what all factors lead to polarization among various social media groups towards a policy (Grover et al. 2018). In case the government receives too much negative opinion towards the policy, then government might consider delaying its implementation or might take expert opinion for its betterment. Once the policy is implemented and if afterwards government starts receiving negative opinion towards the policy as it was in case of GST. Then government should take appropriate steps to overcome the hardship faced by the public.

- (a) **Public/ People:** General public are the other main stakeholder in this system. Public are the one for whom the policy is implemented. As revealed from results of our study, public should engage in social media discussions regarding the new policy (Hossain et al. 2018). Discussion on social media will help the public in understanding the merits as well as demerits of the policy in a better way. They can further raise their concern, or even suggest some recommendation that will indeed help the government for better policy formation of new policy and improvement in the existing policy.
- (b) **Cloud Service Providers:** Cloud service providers such as Amazon web service (AWS - Amazon EC2), Microsoft Azure (2018), Google cloud (Google Cloud 2018) etc. can engage in long term commitment with the government to provide them cloud platform for implementation of such systems. Since it will be a long term commitment, these service providers can gain long term benefits. Further, they can provide additional data analytics services to boost their revenues.

6.3 Limitation and Future Research Directions

Though, the proposed system has given us some encouraging results yet it suffers from certain limitations that needed to be addressed. Firstly, data (Tweets) for experimentation was fetched only from Twitter, while other social media platforms like Facebook, Instagram etc. were not utilized. Secondly, only original tweets were considered for analysis, while no emphasis was given to re-tweets. Re-tweets as we know constitute up to third of the entire Twitter traffic (Holmberg 2014), hence considerable amount of Twitter traffic was ignored by us in our analysis. Similarly, bot detection (Chu et al. 2010) was also not performed in our system. Social media bots can

adversely affect the analysis by making the results biased. Hence, it is very important to address them. Though there are many advantages of cloud computing, still cloud computing has its own limitations such as network communication cost, unreliable latency, lack of mobility support etc. (Stojmenovic and Wen 2014). We can overcome these limitations by extending our system to Fog computing (Yi et al. 2015). All above stated issues can be addressed in future.

7 Conclusions

Technology is advancing with a rapid pace and various governments are also proactive in terms of adopting newer technologies for providing access to enhanced services to its citizens. Multiple efforts are being done to bridge down the perception gap between the policy makers and the general public, for whom these policies are actually designed and implemented. Cloud computing and social media platforms have emerged as two significantly powerful tools for the governments all around the world to communicate and provide quality services to its citizens. However, all the studies conducted till date to the best of our knowledge have utilized the advantages of these tools independently and no fruitful effort has ever been made towards combined implementation of these powerful technologies. We started our research work to combine the selective capabilities of cloud computing and social media analytics towards efficient monitoring and controlling of public policies. We proposed a new cloud based approach, which captures the response of public through data gathered from various social media platforms about new policy implementation. For monitoring of public policies various social media analytics techniques are applied on this captured data (tweets). Based upon the results of these social media analytics techniques, the appropriate controlling operation was performed. We tested our proposed system with data collected towards GST implementation by Indian government.

The initial results indicate a strong support for GST, however as soon as the GST became reality, traders and consumers faced hardship and the overall sentiment dipped towards negative side raising an alarming signal to the government to take appropriate actions. Based on post implementation reaction, government took various controlling measures which again resulted in increase of positive sentiment among citizens. A prominent contribution made in this research was to do location based analysis which can help the government to select the targeted audience affected by a new public policy and can take control measures accordingly. These implementation results were encouraging, suggesting that the system proposed and evaluated in this study can be used for efficient monitoring and controlling of public policies.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abascal-Mena, R., Lema, R., & Sèdes, F. (2015). Detecting sociosemantic communities by applying social network analysis in tweets. *Social Network Analysis and Mining*, 5(1), 38.
- Ahmad, E., & Poddar, S. (2009). *GST reforms and intergovernmental considerations in India*. London: Working paper, 26. Asia research Centre, London School of Economics and Political Science.
- Aladwani, A. M., & Dwivedi, Y. K. (2018). Towards a theory of SocioCitizenry: Quality anticipation, trust configuration, and approved adaptation of governmental social media. *International Journal of Information Management*, 43, 261–272.
- AlAlwan, A., Rana, N. P., Dwivedi, Y. K., & Algharabat, R. (2017). Social Media in Marketing: A review and analysis of the existing literature. *Telematics and Informatics*, 34(7), 1177–1190.
- Alryalat, M. A. A., Rana, N. P., Sahu, G. P., Dwivedi, Y. K., & Tajvidi, M. (2017). Use of social Media in Citizen-Centric Electronic Government Services: A literature analysis. *International Journal of Electronic Government Research (IJEGR)*, 13(3), 55–79.
- Amazon EC2 Instance Price (2018) [Online]. Available: <https://aws.amazon.com/ec2/pricing/on-demand/>
- Amirkhanyan, A., & Meinel, C. (2017). Density and intensity-based spatiotemporal clustering with fixed distance and time radius. In *conference on e-Business, e-Services and e-Society* (pp. 313–324). Cham: Springer.
- Androutsopoulou, A., Charalabidis, Y., & Loukis, E. (2018). Policy informatics in the social media era: Analyzing opinions for policy making. In *International conference on electronic participation* (pp. 129–142). Cham: Springer.
- Arun, R., Suresh, V., Madhavan, C. V., & Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 391–402). Berlin: Springer.
- Attu, R., & Terras, M. (2017). What people study when they study Tumblr: Classifying Tumblr-related academic research. *Journal of Documentation*, 73(3), 528–554.
- AWS - Amazon EC2 (2017) [Online]. Available: <https://aws.amazon.com/ec2/>.
- Bertot, J. C., Jaeger, P. T., & Hansen, D. (2012). The impact of polices on government social media usage: Issues, challenges, and recommendations. *Government Information Quarterly*, 29(1), 30–40.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Bornmann, L., & Haunschild, R. (2016). How to normalize Twitter counts? A first attempt based on journals in the Twitter index. *Scientometrics*, 107(3), 1405–1422.
- Business Standard News (2017a) [Online]. Available: http://www.business-standard.com/article/economy-policy/mrp-retailers-can-sell-gst-inventory-with-new-price-stickers-till-30-sep-117070500333_1.html.
- Business Standard News (2017b) [Online]. Available: http://www.business-standard.com/article/economy-policy/gst-rollout-govt-to-penalise-manufacturers-for-not-reprinting-revised-mrp-117070400756_1.html.
- Buyya, R., Vecchiola, C., & Selvi, S. T. (2013). Mastering cloud computing: Foundations and applications programming. Newnes.

- Ceron, A., & Negri, F. (2016). The “social side” of public policy: Monitoring online public opinion and its mobilization during the policy cycle. *Policy & Internet*, 8(2), 131–147.
- Chae, B. K. (2015). Insights from hashtag# supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics*, 165, 247–259.
- Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2010). Who is tweeting on Twitter: Human, bot, or cyborg?. In proceedings of the 26th annual computer security applications conference (pp. 21–30). ACM.
- Chung, W., & Zeng, D. (2016). Social-media-based public policy informatics: Sentiment and network analyses of US immigration and border security. *Journal of the Association for Information Science and Technology*, 67(7), 1588–1606.
- Clohesy, T., Acton, T., & Morgan, L. (2014). Smart City as a service (SCaaS): A future roadmap for e-government smart city cloud computing initiatives. In *Proceedings of the 2014 IEEE/ACM 7th International conference on utility and cloud computing* (pp. 836–841). IEEE Computer Society.
- Cotterill, S., & King, S. (2007). Public sector partnerships to deliver local e-government: A social network study. In *International conference on electronic government* (pp. 240–251). Berlin: Springer.
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, 1695(5), 1–9.
- Dehak, N., Dehak, R., Glass, J. R., Reynolds, D. A., & Kenny, P. (2010). Cosine similarity scoring without score normalization techniques. In *Odyssey* (p. 15).
- Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique*, 17(1), 61–84.
- Ding, Y. (2011). Community detection: Topological vs. topical. *Journal of Informetrics*, 5(4), 498–514.
- Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. *The Marketing Review*, 15(3), 289–309.
- Dwivedi, Y. K., Shareef, M. A., Simintiras, A. C., Lal, B., & Weerakkody, V. (2016). A generalised adoption model for services: A cross-country comparison of mobile health (m-health). *Government Information Quarterly*, 33(1), 174–187.
- Dwivedi, Y. K., Rana, N. P., Tajvidi, M., Lal, B., Sahu, G. P., & Gupta, A. (2017). Exploring the role of social media in e-government: An analysis of emerging literature. In *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance* (pp. 97–106). ACM.
- Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, R. M. (2017a). An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Information Quarterly*, 34(2), 211–230.
- Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2017b). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-017-9774-y>.
- Economic Times (2017) [Online], Available: <https://economictimes.indiatimes.com/news/economy/policy/gst-rate-guide-know-all-the-things-that-will-get-cheaper-after-midnight/articleshow/59386840.cms>.
- Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.
- Ganis, M., & Kohirkar, A. (2015). Social media analytics: Techniques and insights for extracting business value out of social media. IBM Press.
- García, S., Luengo, J., & Herrera, F. (2016). *Data preprocessing in data mining*. Springer <https://doi.org/10.1007/978-3-319-10247-4>.
- Garima. (2018). From inception to implementation: Evaluating the impact of goods and service tax on Indian E-commerce industry – A critical comment. In S. Moid & S. Dixit (Eds.), *Internet taxation and E-retailing law in the global context* (pp. 123–141). Hershey: IGI Global. <https://doi.org/10.4018/978-1-5225-3787-8.ch008>.
- Google Cloud (2018) [Online]. Available: <https://cloud.google.com/>
- Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2017). The untold story of USA presidential elections in 2016-insights from Twitter analytics. In *Conference on e-Business, e-Services and e-Society* (pp. 339–350). Cham: Springer.
- Grover, P., Kar, A. K., Dwivedi, Y. K., & Janssen, M. (2018). Polarization and acculturation in US election 2016 outcomes—can twitter analytics predict changes in voting preferences. *Technological Forecasting and Social Change*. <https://doi.org/10.1016/j.techfore.2018.09.009>.
- Grubmüller, V., Götsch, K., & Krieger, B. (2013). Social media analytics for future oriented policy making. *European Journal of Futures Research*, 1(1), 1–20.
- GST Benefits (2019) [Online], Available: <http://gstcouncil.gov.in/about-gst>
- Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science*, 17, 26–32.
- Herdağdelen, A., Zuo, W., Gard-Murray, A., & Bar-Yam, Y. (2013). An exploration of social identity: The geography and politics of news-sharing communities in twitter. *Complexity*, 19(2), 10–20.
- Hibbing, J. R., & Theiss-Morse, E. (Eds.). (2001). *What is it about government that Americans dislike?*. Cambridge University Press.
- Holmberg, K. (2014). The impact of retweeting on altmetrics. Retrieved July 8, 2014, from <http://de.slideshare.net/kholmber/the-meaning-of-retweeting>.
- Hossain, M. A., Dwivedi, Y. K., Chan, C., Standing, C., & Olanrewaju, A. S. (2018). Sharing political content in online social media: A planned and unplanned behaviour approach. *Information Systems Frontiers*, 20(3), 485–501.
- IBM (2019) [Online], Available: <https://www.ibm.com/cloud/learn/benefits-of-cloud-computing>
- James, O., & Van Ryzin, G. G. (2017). Incredibly good performance: An experimental study of source and level effects on the credibility of government. *The American Review of Public Administration*, 47(1), 23–35.
- Janssen, M., Rana, N. P., Slade, E. L., & Dwivedi, Y. K. (2018). Trustworthiness of digital government services: Deriving a comprehensive theory through interpretive structural modelling. *Public Management Review*, 20(5), 647–671.
- Jeong, K. H. (2006). E-government, the road to innovation; Principles and Experiences in Korea. Gil-Job-E Media.
- Jha, R. (2018). Regional inequality and indirect tax reform in India. In *Facets of India's economy and her society volume II* (pp. 119–148). London: Palgrave Macmillan.
- Joseph, N., Grover, P., Rao, P. K., & Ilavarasan, P. V. (2017). Deep analyzing public conversations: Insights from Twitter analytics for policy makers. In *Conference on e-Business, e-Services and e-Society* (pp. 239–250). Cham: Springer.
- Kaigo, M., & Okura, S. (2016). Exploring fluctuations in citizen engagement on a local government Facebook page in Japan. *Telematics and Informatics*, 33(2), 584–595.
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present and future. *Information Systems Frontiers*, 20(3), 531–558.
- Kassarjian, H. H. (1977). Content analysis in consumer research. *Journal of Consumer Research*, 4(1), 8–18.
- Kayser, V., & Blind, K. (2017). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*, 116, 208–215.
- Lee, G., & Kwak, Y. H. (2012). An open government maturity model for social media-based public engagement. *Government Information Quarterly*, 29(4), 492–503.
- Liang, D. H., Liang, D. S., & Wen, I. J. (2011). Applications of both cloud computing and e-government in Taiwan. *International Journal of Digital Content Technology and its Applications*, 5(5).

- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on Human Language Technologies*, 5(1), 1–167.
- Liu, Y., Chen, Y., Wu, S., Peng, G., & Lv, B. (2015). Composite leading search index: A preprocessing method of internet search data for stock trends prediction. *Annals of Operations Research*, 234(1), 77–94.
- Llewellyn, C., Grover, C., Alex, B., Oberlander, J., & Tobin, R. (2015). Extracting a topic specific dataset from a Twitter archive. In *International Conference on Theory and Practice of Digital Libraries* (pp. 364–367). Cham: Springer.
- McNaught, C., & Lam, P. (2010). Using Wordle as a supplementary research tool. *The Qualitative Report*, 15(3), 630–643.
- Microsoft Azure (2018) [Online], Available: <https://azure.microsoft.com/en-in/>
- Mishra, N., & Singh, A. (2016). Use of twitter data for waste minimisation in beef supply chain. *Annals of Operations Research*, 270(1–2), 1–23. <https://doi.org/10.1007/s10479-016-2303-4>.
- Mohammad, Saif M., and Peter D. Turney. (2010) Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, pp. 26–34. Association for Computational Linguistics.
- Mohammed, F., Ibrahim, O., & Ithnin, N. (2016). Factors influencing cloud computing adoption for e-government implementation in developing countries: Instrument development. *Journal of Systems and Information Technology*, 18(3), 297–327.
- Mossberger, K., Wu, Y., & Crawford, J. (2013). Connecting citizens and local governments? Social media and interactivity in major US cities. *Government Information Quarterly*, 30(4), 351–358.
- NDTV News (2017) [Online], Available: <https://www.ndtv.com/education/100-hour-certificate-course-in-gst-launched-1717191>.
- Ou, Gaoyan, Wei Chen, Tengjiao Wang, Zhongyu Wei, L. I. Binyang, Dongqing Yang, and Kam-Fai Wong. (2014) Exploiting community emotion for microblog event detection. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1159–1168.
- Park, M. J., Kang, D., Rho, J. J., & Lee, D. H. (2016). Policy role of social media in developing public trust: Twitter communication with government leaders. *Public Management Review*, 18(9), 1265–1288.
- Pfeffermann, D., Eltinge, J. L., Brown, L. D., & Pfeffermann, D. (2015). Methodological issues and challenges in the production of official statistics: 24th annual Morris Hansen lecture. *Journal of Survey Statistics and Methodology*, 3(4), 425–483.
- Pieterse, W., & Johnson, Z. (2011). Multi-Channel management: Recent developments in PES and E-Government: Analytical Paper. PES to PES dialogue, The European Commission Mutual Learning Programme for Public Employment Services.
- Purohit, H., Hampton, A., Shalin, V. L., Sheth, A. P., Flach, J., & Bhatt, S. (2013). What kind of conversation is Twitter? Mining psycholinguistic cues for emergency coordination. *Computers in Human Behavior*, 29(6), 2438–2447.
- Rana, N. P., Dwivedi, Y. K., & Williams, M. D. (2013). Analysing challenges, barriers and CSF of egov adoption. *Transforming Government: People, Process and Policy*, 7(2), 177–198.
- Rana, N. P., Dwivedi, Y. K., & Williams, M. D. (2015). A meta-analysis of existing research on citizen adoption of e-government. *Information Systems Frontiers*, 17(3), 547–563.
- Rana, N. P., Dwivedi, Y. K., Williams, M. D., & Weerakkody, V. (2016). Adoption of online public grievance redressal system in India: Toward developing a unified view. *Computers in Human Behavior*, 59, 265–282.
- Rana, N. P., Dwivedi, Y. K., Lal, B., Williams, M. D., & Clement, M. (2017). Citizens' adoption of an electronic government system: Towards a unified view. *Information Systems Frontiers*, 19(3), 549–568.
- Rosenstone, S. J., Hansen, J. M., & Reeves, K. (1993). *Mobilization, participation, and democracy in America*. New York: Macmillan Publishing Company.
- Saboo, A. R., Kumar, V., & Park, I. (2016). Using big data to model time-varying effects for marketing resource (re) allocation. *MIS Quarterly*, 40(4), 911–939.
- Sadiku, M. N., Musa, S. M., & Momoh, O. D. (2014). Cloud computing: Opportunities and challenges. *IEEE Potentials*, 33(1), 34–36.
- Saif, H., Fernandez, M., He, Y., & Alani, H. (2013). Evaluation datasets for Twitter sentiment analysis: A survey and a new dataset, the STS-gold.
- Schneider, S. K., Jacoby, W. G., & Lewis, D. C. (2010). Public opinion toward intergovernmental policy responsibilities. *Publius: The Journal of Federalism*, 41(1), 1–30.
- Severo, M., Feredj, A., & Romele, A. (2016). Soft data and public policy: Can social media offer alternatives to official statistics in urban policymaking? *Policy & Internet*, 8(3), 354–372.
- Shareef, M. A., Dwivedi, Y. K., Kumar, V., & Kumar, U. (2016). Reformation of public service to meet citizens' needs as customers: Evaluating SMS as an alternative service delivery channel. *Computers in Human Behavior*, 61, 255–270.
- Sharma, R., Sharma, A., & Singh, R. R. (2012). E-governance & cloud computing: Technology oriented government policies. *International Journal of Research in IT & Management*, 2(2), 584–593.
- Shiau, W.-L., Dwivedi, Y. K., & Yang, H.-S. (2017). Co-citation and cluster analyses of extant literature on social networks. *International Journal of Information Management*, 37(5), 390–399.
- Shiau, W.-L., Dwivedi, Y. K., & Lai, H.-H. (2018). Examining the core knowledge on Facebook. *International Journal of Information Management*, 43, 52–63.
- Singh, P., Sawhney, R. S., & Kahlon, K. S. (2017). Forecasting the 2016 US presidential elections using sentiment analysis. In *Conference on e-Business, e-Services and e-Society* (pp. 412–423). Cham: Springer.
- Singh, P., Dwivedi, Y. K., Kahlon, K. S., & Sawhney, R. S. (2018a). Intelligent monitoring and controlling of public policies using social media and cloud computing. In *International working conference on transfer and diffusion of IT* (pp. 143–154). Cham: Springer.
- Singh, P., Sawhney, R. S., & Kahlon, K. S. (2018b). Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. *ICT Express*, 4(3), 124–129.
- Singh, P., Kahlon, K. S., Sawhney, R. S., Vohra, R., & Kaur, S. (2018c). Social media buzz created by # nanotechnology: Insights from Twitter analytics. *Nanotechnology Reviews*, 7(6), 521–528.
- Singh, P., Sawhney, R. S., & Kahlon, K. S. (2019). Twitter based sentiment analysis of GST implementation by Indian government. In *Digital business* (pp. 409–427). Cham: Springer.
- Smitha, K. K., Thomas, T., & Chitharanjan, K. (2012). Cloud based e-governance system: A survey. *Procedia Engineering*, 38, 3816–3823.
- Stephens, M., & Poorthuis, A. (2015). Follow thy neighbor: Connecting the social and the spatial networks on Twitter. *Computers, Environment and Urban Systems*, 53, 87–95.
- Stieglitz, S., & Dang-Xuan, L. (2013a). Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217–248.
- Stieglitz, S., & Dang-Xuan, L. (2013b). Social media and political communication: A social media analytics framework. *Social Network Analysis and Mining*, 3(4), 1277–1291.
- Stojmenovic, I., & Wen, S. (2014). The fog computing paradigm: Scenarios and security issues. In *Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on* (pp. 1–8). IEEE.
- The Constitution (One Hundred and First Amendment) Act (2017) [Online]. Available: [http://lawmin.nic.in/ld/The%20Constitution%20\(One%20Hundred%20and%20First%20Amendment\)%20Act,%202016.pdf](http://lawmin.nic.in/ld/The%20Constitution%20(One%20Hundred%20and%20First%20Amendment)%20Act,%202016.pdf).
- Twitter API Price (2018) [Online]. Available: <https://developer.twitter.com/en/pricing.html>

- Van den Broeck, J., Cunningham, S. A., Eeckels, R., & Herbst, K. (2005). Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Medicine*, 2, no. 10, e267.
- Walther, M., & Kaissar, M. (2013). Geo-spatial event detection in the twitter stream. In *European conference on information retrieval* (pp. 356–367). Berlin: Springer.
- Yi, S., Li, C., & Li, Q. (2015). A survey of fog computing: Concepts, applications and issues. In *Proceedings of the 2015 workshop on mobile big data* (pp. 37–42). ACM.
- Yuan, H., Xu, W., Li, Q., & Lau, R. (2017). Topic sentiment mining for sales performance prediction in e-commerce. *Annals of Operations Research*, 270(1–2), 1–24. <https://doi.org/10.1007/s10479-017-2421-7>.
- Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-scale network analysis for online social brand advertising. *MIS Quarterly*, 40(4), 849–868.
- Zwattendorfer, B., Stranacher, K., Tauber, A., & Reichstädter, P. (2013). Cloud computing in e-government across europe. In *International conference on electronic government and the information systems perspective* (pp. 181–195). Berlin: Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Prabhsimran Singh is an Assistant Professor in Department of Computer Engineering & Technology, Guru Nanak Dev University, Amritsar, India. He has a teaching experience of 4 years. He holds a Master of Technology from Guru Nanak Dev University, Amritsar. His area of interest includes Social Media Analytics, Text Mining and Data Security.

Yogesh K. Dwivedi is a Professor of Digital Marketing and Innovation, the University Dean of Academic Leadership (REF Research Environment), Founding Director of the Emerging Markets Research Centre (EMaRC) and Co-Director of Research at the School of Management, Swansea University, Wales, UK. Professor Dwivedi is also currently leading the *International Journal of Information Management* as its Editor-in-Chief. His research interests are at the interface of Information Systems (IS) and Marketing, focusing on issues related to consumer adoption and diffusion of emerging digital innovations, digital government, and digital marketing particularly in the context of emerging markets. Professor Dwivedi has published more than 300 articles in a range of leading academic journals and conferences that are widely cited (more than 10 thousand times as per Google Scholar). Professor Dwivedi is an Associate Editor of the *European Journal of Marketing*, *Government Information Quarterly* and *International Journal of Electronic Government Research*, and Senior Editor of the *Journal of Electronic Commerce Research*. More information about Professor Dwivedi

can be found at: <http://www.swansea.ac.uk/staff/som/academic-staff/y.k.dwivedi/>

Karanjeet Singh Kahlon is a Professor at Department of Computer Science, Guru Nanak Dev University, Amritsar & Registrar, Guru Nanak Dev University, Amritsar, India. He has done his PhD in Computer Science & Engineering. He has more than 23 years of Teaching and Research Experience. His area of interest includes Networking and Distributed Computing.

Ravinder Singh Sawhney is Associate Professor and Head at Department of Electronics Technology, Guru Nanak Dev University, Amritsar, India. He has done his PhD in the field of Nano electronics and has worked on modeling and simulation of charge transport through various molecules and designing of various molecular junctions for future Nano electronic Devices. He has teaching experience of more than 20 years to both post-graduate and under graduate engineering students. He has more than 75 publications to his credit in various International journals. He has membership of many International and National Engineering Societies. He also has interest in the field of Data Communication, especially WLANs, MANETS, Mesh Networks, MPLS and Social Media Analytics. Currently, he is working Bio-Electric Implants in collaboration with IIT, Ropar. He has been on the board of many journals as reviewer.

Ali Abdallah Alalwan is an Assistant Professor at Amman Collage of Banking and Finance at Al-Balqa Applied University, Jordan. He holds a Bachelor's Degree in Marketing and an MBA/Marketing degree from the University of Jordan. He also holds a PhD from Swansea University. His current research interest is in the area of Information Systems, Technology acceptance, electronic marketing, social media, Internet of things, self-service technologies, Internet banking, and Mobile banking. A part of his work has been published in some refereed journals including JFSM, JEIM, ISM, IJIM.

Nripendra P Rana is Professor in Information Systems and the Director of Post Graduate Research (PGR) in the School of Management at Swansea University, UK. With an academic and professional background in Mathematics and Computer Science, his current research interests focus on the adoption of emerging and disruptive technologies, digital and social media marketing, Internet of Things (IoT) and smart cities, Open Data and the use of ICT for Learning and Teaching in Higher Education. He is an experienced quantitative researcher and his work has been published in leading academic journals such as *European Journal of Marketing*, *Journal of Business Research*, *Information Systems Frontiers*, *Government Information Quarterly*, *Computers in Human Behavior*, *Annals of Operations Research*, *Technological Forecasting and Social Change*, *International Journal of Production Research*, *International Journal of Production Economics*, *Information Technology & People*, *Public Management Review*, *Production Planning & Control*, and *International Journal of Information Management*.